# 7

# Answers and Hints

### 7.2 Properties of Evolutionary Stable Strategies

(a) Simple

(b) The two conditions for an evolutionarily stable strategy imply that if $\sigma$ is an evolutionarily stable strategy, then $\pi_{\tau\sigma} \leq \pi_{\sigma\sigma}$ for any other strategy $\tau$. But this is just the Nash condition.

(c) If $\sigma$ is a strict Nash equilibrium, then $\pi_{\tau\sigma} < \pi_{\sigma\sigma}$ for any $\tau \neq \sigma$, so $\sigma$ is an evolutionarily stable strategy.

(d) Suppose $\pi_{11} > \pi_{21}$. Then pure strategy 1 is a strict Nash equilibrium, so it is an evolutionarily stable strategy. Similarly if $\pi_{22} > \pi_{12}$. So suppose $\pi_{11} < \pi_{21}$ and $\pi_{22} < \pi_{12}$. Then we can show that the game has a unique completely mixed symmetric equilibrium $p$, where each player uses strategy 1 with probability $\alpha_p \in (0, 1)$. The payoff to strategy 1 against the mixed strategy $(\alpha_p, 1 - \alpha_p)$ is then $\alpha_p\pi_{11} + (1 - \alpha_p)\pi_{12}$ and the payoff to strategy 2 against this mixed strategy is $\alpha_p\pi_{21} + (1 - \alpha_p)\pi_{22}$. Since these must be equal, we find that $\alpha_p = (\pi_{22} - \pi_{12})/\Delta$, where $\Delta = \pi_{11} - \pi_{21} + \pi_{22} - \pi_{12} < 0$. Notice that under our assumptions, $0 < \alpha_p < 1$, so there is a unique completely mixed Nash equilibrium $(\alpha_p, 1 - \alpha_p)$. Is this evolutionarily stable?

Let $\alpha_q$ be the probability a mutant player uses pure strategy 1. Since each pure strategy is a best response to $\alpha_p$, $\alpha_q$ must also be a best evolutionarily stable, we must show that $\pi_{pq} > \pi_{pq}$. We have

$$\pi_{pq} = \alpha_p[a_{11}\alpha_q + \pi_{12}(1 - \alpha_q)] + (1 - \alpha_p)[\pi_{21}\alpha_q + \pi_{22}(1 - \alpha_q)]$$

and

$$\pi_{qq} = \alpha_q[\pi_{11}\alpha_q + \pi_{12}(1 - \alpha_q)] + (1 - \alpha_q)[\pi_{21}\alpha_q + \pi_{22}(1 - \alpha_q)].$$

VISIT…

Subtracting and simplifying, we get

$$\pi_{pq} - \pi_{qq} = -(\alpha_p - \alpha_q)^2\Delta > 0,$$

which proves evolutionary stability.

(e) It is easy to check that if there is a mixed strategy equilibrium, the frequency $\alpha$ of pure strategy 1 must satisfy

$$\alpha = \frac{\pi_{22} - \pi_{12}}{\Delta} \qquad \text{where } \Delta = \pi_{11} - \pi_{21} + \pi_{22} - \pi_{12}.$$

Suppose $\Delta > 0$. Then $0 < \alpha < 1$ if and only if $0 < \pi_{22} - \pi_{12} < \pi_{11} - \pi_{21} + \pi_{22} - \pi_{12}$, which is true if and only if $\pi_{11} > \pi_{21}$ and $\pi_{22} > \pi_{12}$. If $\Delta < 0$, a similar argument shows that $0 < \alpha < 1$ if and only if the other pair of inequalities holds.

Suppose there is a 'mutant' that uses pure strategy 1 with probability $\beta$. Thus in general,

$$\begin{aligned}\pi_{\gamma\delta} &= \gamma\delta\pi_{11} + \gamma(1-\delta)\pi_{12} + (1-\gamma)\delta\pi_{21} + (1-\gamma)(1-\delta)\pi_{22} \\ &= \gamma\delta\Delta + \delta(\pi_{21} - \pi_{22}) + \gamma(\pi_{12} - \pi_{22}) + \pi_{22}.\end{aligned}$$

It follows that

$$\pi_{\alpha\alpha} - \pi_{\beta\alpha} = (\alpha - \beta)[\alpha\Delta - (\pi_{22} - a_{12})] = 0,$$

so the equilibrium is evolutionarily stable if and only if $\pi_{\alpha\beta} > \pi_{\beta\beta}$. But

$$\begin{aligned}\pi_{\alpha\beta} - \pi_{\beta\beta} &= \alpha\beta\Delta + \beta(a_{21} - a_{22}) + \alpha(a_{12} - a_{22}) + a_{22} \\ &\qquad - \beta^2\Delta - \beta(a_{21} - a_{22}) - \beta(a_{12} - a_{22}) - a_{22} \\ &= \beta(\alpha - \beta)\Delta + (\alpha - \beta)(a_{12} - a_{22}) \\ &= (\alpha - \beta)(\beta\Delta + a_{12} - a_{22}) \\ &= (\alpha - \beta)(\beta\Delta - \alpha\Delta) \\ &= -(\alpha - \beta)^2\Delta.\end{aligned}$$

Thus the equilibrium is evolutionarily stable if and only if $\Delta < 0$, which is equivalent to $a_{11} < a_{21}$ and $a_{22} < a_{12}$. This proves the assertion.

(f) Suppose there are an infinite number of distinct evolutionarily stable strategies. Then there must be two, say $\sigma$ and $\tau$, that use exactly

the same pure strategies. Now $\tau$ is a best response to $\sigma$, so $\sigma$ must do better against $\tau$ than $\tau$ does against itself. But $\sigma$ does equally well against $\tau$ as $\tau$ does against $\tau$. Thus $\sigma$ is not evolutionarily stable, and similarly for $\tau$.

(g) First, suppose $\sigma$ is evolutionarily stable, so for any $\tau \neq \sigma$, there is an $\tilde{\epsilon}(\tau)$ such that[1]

$$\pi_{\tau,(1-\epsilon)\sigma+\epsilon\tau} < \pi_{\sigma,(1-\epsilon)\sigma+\epsilon\tau} \quad \text{for all } \epsilon \in (0, \tilde{\epsilon}(\tau)). \tag{7.1}$$

In fact, we can choose $\tilde{\epsilon}(\tau)$ as follows. If (7.1) holds for all $\epsilon \in (0,1)$, then let $\tilde{\epsilon}(\tau) = 1$. Otherwise let $\tilde{\epsilon}$ be the smallest $\epsilon > 0$ such that (7.1) is violated, and define

$$\tilde{\epsilon}(\tau) = \frac{\pi_{\sigma\sigma} - \pi_{\tau\sigma}}{\pi_{\tau\tau} - \pi_{\tau\sigma} - \pi_{\sigma,\tau} + \pi_{\sigma\sigma}}.$$

It is easy to check that $\tilde{\epsilon}(\tau) \in (0,1]$ and (7.1) is satisfied. Let $T \subset S$ be the set of strategies such that if $\tau \in T$, then there is at least one pure strategy used in $\sigma$ that is not used in $\tau$. Clearly $T$ is closed and bounded, $\sigma \notin T$, $\tilde{\epsilon}(\tau)$ is continuous, and $\tilde{\epsilon}(\tau) > 0$ for all $\tau \in T$. Hence $\tilde{\epsilon}(\tau)$ has a strictly positive minimum $\epsilon^*$ such that (7.1) holds for all $\tau \in T$ and all $\epsilon \in (0, \epsilon^*)$.

If $\tau$ is a mixed strategy, and $s$ is a pure strategy, we define $s(\tau)$ to be the weight of $s$ in $\tau$ (i.e., the probability that $s$ will be played using $\tau$). Now consider the neighborhood of $s$ consisting of all strategies $\tau$ such that $|1 - s(\tau)| < \epsilon^*$ for all pure strategies $s$. If $\tau \neq s$, then $\epsilon^* > 1 - s(\tau) = \epsilon > 0$ for some pure strategy $s$. Then $\tau = (1-\epsilon)s + \epsilon r$ where $r \in T$. But then (7.1) gives $\pi_{r\tau} < \pi_{s,\tau}$. If we multiply both sides of this inequality by $\epsilon$ and add $(1-\epsilon)\pi_{s\tau}$ to both sides, we get $\pi_{\tau\tau} < \pi_{s,\tau}$, as required. The other direction is similar, which proves the assertion.

(h) If $\sigma$ is completely mixed, then for any $t \in S$, $\pi_{\sigma\sigma} = \pi_{t\sigma}$, simply because any pure strategy has the same payoff against $\sigma$ as $\sigma$ does against $\sigma$. Therefore any mixed strategy has the same payoff against $\sigma$ as $\sigma$ has against $\sigma$. For similar reasons $\pi_{\sigma,\tau} = \pi_{\sigma\sigma}$. Thus $\sigma$ is evolutionarily stable and $\tau$ is any other strategy, we must have $\pi_{\sigma\tau} > \pi_{\tau\tau}$.

[1]We write $(a,b)$ to mean the set of numbers $\{x | a < x < b\}$, and we call this the *open interval* $(a,b)$.

### 7.4 Cooperative Fishing

Here is the normal form game:

| | Put Out | Free Ride |
|---|---|---|
| Put Out | $\frac{\nu}{2}-c_2, \frac{\nu}{2}-c_2$ | $\frac{\nu}{2}-c_1, \frac{\nu}{2}$ |
| Free Ride | $\frac{\nu}{2}, \frac{\nu}{2}-c_1$ | 0,0 |

It is easy to see there are no pure strategy symmetric equilibria, since $\nu/2 > c_1$. There are two pure strategy asymmetric equilibria, $FP$ and $PF$. Consider a mixed strategy equilibrium where a fraction $\alpha$ of the population play $P$. The payoff to $P$ is then

$$\alpha\left(\frac{\nu}{2}-c_2\right)+(1-\alpha)\left(\frac{\nu}{2}-c_1\right)=\frac{\nu}{2}-[\alpha c_2+(1-\alpha)c_1].$$

The payoff to $F$ is simply $\alpha\nu/2$. Equating the two payoffs, we get

$$\alpha=\frac{\frac{\nu}{2}-c_1}{\frac{\nu}{2}+c_2-c_1}.$$

Notice that we have $0<\alpha<1$, so this is a strictly mixed Nash equilibrium. Is this mixed strategy, which we will call $M$, an evolutionarily stable strategy?

Let $\pi_{AB}$ be the payoff to the agent playing $A$ when the other agent plays $B$. Let $N \neq M$, so clearly $\pi_{NM}=\pi_{MM}$, since both pure strategies in $N$ have equal payoffs against $M$ (or else $M$ would not be Nash, since $M$ uses the same pure strategies as $N$). If this is *not* clear, consider the following argument: Suppose strategy $N$ uses Put Out with probability $\beta$. Then $\pi_{NM}=\beta\pi_{PM}+(1-\beta)\pi_{FM}=\beta\pi_{MM}+(1-\beta)\pi_{MM}=\pi_{MM}$, the next-to-last step following from the fact that $M$ is a mixed strategy Nash equilibrium, so both $P$ and $F$ have equal payoffs against $M$.

We thus must show that $\pi_{MN}>\pi_{NN}$ to prove that $M$ is evolutionarily stable. Define $\pi_{\gamma\delta}$ as the payoff to player 1 when 1 uses Put Out with probability $\gamma$ and 2 uses Put Out with probability $\delta$. We have

$$\begin{aligned}\pi_{\gamma\delta}&=\gamma\left[\delta\left(\frac{\nu}{2}-c_2\right)+(1-\delta)\left(\frac{\nu}{2}-c_1\right)\right]+(1-\gamma)\delta\frac{\nu}{2}\\&=\frac{\nu}{2}[\gamma+(1-\gamma)\delta]+\gamma\delta(c_1-c_2)-\gamma c_1.\end{aligned}$$

But $\pi_{MN} = \pi_{\alpha\beta}$ and $\pi_{NN} = \pi_{\beta\beta}$, so

$$\begin{aligned}
\pi_{MN} - \pi_{NN} &= \pi_{\alpha\beta} - \pi_{\beta\beta} \\
&= \frac{\nu}{2}[\alpha + (1-\alpha)\beta] + \alpha\beta(c_1 - c_2) - \alpha c_1 \\
&\qquad - \frac{\nu}{2}[\beta + (1-\beta)\beta] - \beta^2(c_1 - c_2) + \beta c_1 \\
&= \frac{\nu}{2}[(\alpha - \beta) + (\beta - \alpha)\beta] \\
&\qquad + (\alpha - b)\beta(c_1 - c_2) - (\alpha - \beta)c_1 \\
&= (\alpha - \beta)\left[\frac{\nu}{2}(1-\beta) + \beta(c_1 - c_2) - c_1\right] \\
&= (\alpha - \beta)\left[\frac{\nu}{2} - c_1 - \beta\left[\frac{\nu}{2} - c_1 + c_2\right]\right] \\
&= (\alpha - \beta)^2\left[\frac{\nu}{2} - c_1 + c_2\right] > 0.
\end{aligned}$$

This completes the proof that the equilibrium is evolutionarily stable.

### 7.11 Hawks, Doves and Bourgeois

(a) Easy

(b) Piece o' cake

(c) The payoff to H is $\alpha(v-w)/2 + (1-\alpha)v = v - \alpha(v+w)/2$, and the payoff to D is $(1-\alpha)(v/2 - t) = v/2 - t - \alpha(v/2 - t)$. These are equated when $\alpha = (v+2t)/(w+2t)$, which is $< 1$ if $w > v$ and $t > 0$.

To show that this mixed strategy equilibrium is evolutionarily stable, we can refer to the 'Theorem' proved earlier, which says that

evolutionary stability is equivalent to $a_{HH} < a_{HD}$, which is true. Or we can do it out in full (it's going to look a lot like the proof of the 'Theorem'). Here is the proof in full:

Suppose player 2 plays $\alpha H + (1-\alpha)D$ and player 1 uses $\beta H + (1-\beta)D$. Then

$$\begin{aligned}\pi_{\beta\alpha} &= \alpha\beta\frac{v-2}{2} + \beta(1-\alpha)v + (1-\alpha)(1-\beta)\left[\frac{v}{2} - t\right] \\ &= -\alpha\beta\left[\frac{w}{2} + t\right] - \alpha\left[\frac{v}{2} - t\right] + \beta\left[\frac{v}{2} + t\right] + \frac{v}{2} - t.\end{aligned}$$

Now for evolutionary stability we must have either $\pi_{\alpha\alpha} > \pi_{\beta\alpha}$, or $\pi_{\alpha\alpha} = \pi_{\beta\alpha}$ and $\pi_{\alpha\beta} > \pi_{\beta\beta}$. But

$$\begin{aligned}\pi_{\alpha\alpha} - \pi_{\beta\alpha} &= -\alpha^2\left[\frac{w}{2} + t\right] + 2\alpha t + \frac{v}{2} - t - \pi_{\beta\alpha} \\ &= -\alpha(\alpha-\beta)\left[\frac{w}{2} + t\right] + \alpha\left[\frac{v}{2} + t\right] - \beta\left[\frac{v}{2} + t\right] \\ &= (\alpha-\beta)\left[-\alpha\left[\frac{w}{2} + t\right] + \frac{v}{2} + t\right] = 0.\end{aligned}$$

Also

$$\pi_{\alpha\beta} - \pi_{\beta\beta} = -(\beta-\alpha)\left[-\beta\left[\frac{w}{2} + t\right] + \frac{v}{2} + t\right].$$

The term in brackets is positive if $\beta < \alpha$ and negative if $\beta > \alpha$, which proves evolutionary stability.

(d) $\pi_{BB} > \pi_{HB}$ if and only if $v/2 > 3v/4 - w/4$, which is equivalent to $w > v$. $\pi_{BB} = \pi_{HB}$ and $\pi_{BH} > \pi_{HH}$ if and only if $v = w$ and $(v-w)/4 > (v-w)/2$, which is false. Thus $B$ can't be invaded by $H$ if $w > v$. Meanwhile $\pi_{BB} > \pi_{DB}$ if and only if $v/2 > v/4 - t/2$, which is true.

### 7.12 Trogs and Farfel

(a) The payoff matrix to the game is:

| | Good | Bad |
|---|---|---|
| Good | $z-y, z-y$ | $w-y, z-x$ |
| Bad | $z-x, w-y$ | $w-x, w-x$ |

By definition, a $2 \times 2$ prisoner's dilemma is given by

| | C | D |
|---|---|---|
| C | $r, r$ | $s, t$ |
| D | $t, s$ | $p, p$ |

where $t > r > p > s$ and $r > (s + t)/2$. This is true if and only if $z - w > y - x$, as can be seen as follows. In our case $t > r > p > s$ becomes $z - x > z - y > w - x > w - y$. The first and third inequalities are always true, and the second is equivalent to $z - w > y - x$. Also $r > (s + t)/2$ is equivalent to $2z - 2y > z - x + w - y$, which is equivalent to $z - y > w - x$.

(b) Obvious, since "always give good farfel" leads to cooperation on every round, either against itself or against Tit-for-Tat.

(c) First we show that playing B forever is better than Tit-for-Tat if and only if $\delta < (y - x)/(z - w)$. Tit-for-Tat against Tit-for-Tat gives $z - y$ for all rounds, with present value $(z - y)/(1 - \delta)$. The present value of defecting forever is $(z - x) + \delta(w - x)/(1 - \delta)$. But $(z - y)/(1 - \delta) > (z - x) + \delta(w - x)/(1 - \delta)$ if and only if $z - y > (1 - \delta)(z - x) + \delta(w - x) = z - x - \delta z + \delta w$ if and only if $y - x < \delta(z - w)$ if and only if $\delta > (y - x)/(z - w)$.

Next we show that playing B for a certain number of periods, say $n \geq 1$, followed by a G is not a best response to Tit-for-Tat if and only if the above inequality holds. The history of play then looks like

$$GGG...BB.....BGxxxxx...$$
$$GGG...GB.....BBGxxxx...,$$

where we don't know what the $x$'s are. Let us compare this to what happens if player 1 defects one fewer time, so we now have

$$GGG...BB.....GGxxxxx...$$
$$GGG...GB.....BGGxxxx....$$

The only changes in the payoff to player 1 are in periods $n - 1$ and $n$. Player 1 gains $(w - y) - (w - x) = x - y$ (which is negative)

in period $n-1$ and gains $(z-y)-(w-y)=z-w$ in period $n$. The total gain, discounted to period $n-1$, is then $x-y+\delta(z-w)$ which is positive if and only if the above inequality holds.

### 7.13 Evolutionary Stability in Finite Populations

(a) Let $r_{\mu\nu}=0$, $r_{\nu\mu}=n$, $r_{\mu\mu}=n+1$, and $r_{\nu\nu}=-1$. Then $\nu$ is not Nash, since $r_{\mu\nu}>r_{\nu\nu}$, $\mu$ is Nash since $r_{\mu\mu}>r_{\mu\nu}$ and $r_{\mu\mu}>r_{\nu\mu}$, but $r(\nu)-r(\mu)=1/n>0$ for any $m$.

(b) This is trivial if there is an alternate best response to a Nash strategy.

(c) Suppose $\nu$ is evolutionarily stable but is not Nash. Then there is some $\mu$ such that $r_{\nu\nu}<r_{\mu\nu}$. Let $m=1$. Then for sufficiently large $n$ we have $r(\nu)<r(\mu)$ in

$$r(\nu)-r(\mu)=\left(1-\frac{m}{n}\right)(r_{\nu\nu}-r_{\mu\nu})+\frac{m}{n}(r_{\nu\mu}-r_{\mu\mu})+\frac{1}{n}(r_{\mu\mu}-r_{\mu\nu}). \tag{7.2}$$

Hence $\nu$ must be Nash. Now suppose $\nu$ is evolutionarily stable and $r_{\nu\nu}=r_{\mu\nu}$ but $r_{\nu\mu}<r_{\mu\mu}$. Equation (7.2) becomes

$$r(\nu)-r(\mu)=\frac{1}{n}\left\{m[r_{\nu\mu}-r_{\mu\mu}]+[r_{\mu\mu}-r_{\mu\nu}]\right\}.$$

Given $\epsilon>0$, choose $\overline{m}$ so that the term in brackets is negative, and then choose $n$ so that $\overline{m}/n<\epsilon$. Then $r(\nu)<r(\mu)$ for all positive $m\leq\overline{m}$, which is a contradiction. So suppose in addition to $r_{\nu\nu}=r_{\mu\nu}$ and $r_{\nu\mu}=r_{\mu\mu}$, we have $r_{\mu\mu}<r_{\mu\nu}$. Then clearly $r(\nu)-r(\mu)=[r_{\mu\mu}-r_{\mu\nu}]/n<0$, again a contradiction. This proves that the stated conditions are necessary. We can reverse the argument to prove the conditions are sufficient as well.

(d) In the limit we have

$$r(\nu)-r(\mu)=(1-\epsilon)[r_{\nu\nu}-r_{\mu\nu}]+\epsilon[r_{\nu\mu}-r_{\mu\mu}].$$

The conclusion follows immediately from this equation. The limit argument cannot be use to conclude that $r(\nu)>r(\mu)$ in the 'large finite' case if $r_{\nu\nu}=r_{\mu\nu}$ and $r_{\nu\mu}=r_{\mu\mu}$.